DOCUMENT RESUME

ED 344 891 TM 017 970

AUTHOR	Campbell, Kathleen Taylor; Tucker, Mary L.
TITLE	The Use of Commonality Analysis in Multivariate Canonical Correlation Analysis.
PUB DATE	Feb 92
NOTE	19p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 31-February 2, 1992).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Assistant Principals; *Correlation; Elementary Secondary Education; *Multivariate Analysis; Personality Measures; *Predictor Variables; Principals; *Raw Scores; *Regression (Statistics)
IDENTIFIERS	*Commonality Analysis; Criterion Variables; Dependent Variables; Linear Models; Myers Briggs Type Indicator; Parametric Analysis

ABSTRACT

Since canonical correlation analysis subsumes multiple regression as a special case, and since commonality analysis (a variance partitioning procedure) has proven useful in interpreting multiple regression results, the interpretation of canonical correlation results might also be enhanced by the use of commonality analysis. In this paper, a canonical correlation analysis using a data set of 64 observations, consisting of four predictor (independent) variables and four criterion (dependent) variables, is interpreted. Predictor variables were raw scores on four sections of the Myers-Briggs Type Indicator. Subjects were 64 school principals and assistant principals. A commonality analysis of the data is explained and illustrated, and results are interpreted in conjunction with the canonical correlation analyses. Commonality analysis offers several advantages for interpretation in that it: (1) indicates the degree to which predictors share variance with criterion variables; (2) indicates the extent of overlap of the variables; and (3) reinforces the recognition that canonical correlation is the most general linear model of parametric statistics. Four tables present analysis results, and a 20-item list of references is included. (SLD)

THE USE OF COMMONALITY ANALYSIS IN MULTIVARIATE CANONICAL CORRELATION ANALYSIS

Kathleen Taylor Campbell, Ph.D.
University of New Orleans

Mary L. Tucker, Ph.D.
Nicholls State University

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX, February 1, 1992.

THE USE OF COMMONALITY ANALYSIS IN MULTIVARIATE CANONICAL CORRELATION ANALYSIS

The paper illustrates how a commonality analysis is performed and used in the interpretation of a canonical correlation analysis. A small data set from a larger study is used to make the discussion concrete and to enable readers to conduct their own multivariate commonality analysis.

THE USE OF COMMONALITY ANALYSIS IN MULTIVARIATE CANONICAL CORRELATION ANALYSIS

The importance of using multivariate statistics in behavioral studies has been pointed out in recent years by researchers (Fish, 1988; Huberty and Morris, 1989; LeCluyse, 1990; Thompson, 1986b). As Fish (1988) states, the term multivariate is used by social science researchers to describe "those statistical techniques that examine two or more dependent variables simultaneously" (p. 130), and the most frequently used multivariate methods are multivariate analysis of variance (MANOVA), discriminant analysis, and canonical correlation analysis. LeCluyse (1990) maintains that since human behavior has multiple facets which are each affected by a wide range of variables, "many behavioral studies ask questions that involve both multiple independent and multiple dependent variables" (p. 1). The use of univariate methods in those instances, she further contends, is inappropriate.

Fish (1988) points out three reasons why multivariate statistics are usually vital in social science research. The first reason is that multivariate methods are used to control the experimentwise Type I error rates. In order to avoid Type I errors, researchers often set their alpha levels quite low; however, what some researchers may not realize is that when several hypotheses are tested on the same sample of subjects (as in the case of multiple univariate tests), the experimentwise alpha level is inflated to an unknown degree. The use of multivariate statistics, however, which allows the testing of many hypotheses simultaneously, controls such inflation, and helps prevent the researcher from committing a Type I error of erroneously rejecting the null hypothesis.

The second reason for using multivariate statistical techniques is to detect statistically significant results which may go unnoticed with the use of univariate statistics. Although statistical significance is mainly influenced by sample size, and it is "not the end-all and be-all of research" (Fish, p. 131), it is still the benchmark by which many journal editors guide their publication decisions. It should, therefore, be noteworthy that the "failure to employ multivariate methods can lead to a failure to identify statistically significant results which actually exist" (Thompson, 1988b, p. 12). This failure has been illustrated by Fish (1988).

The third and most important reason for using multivariate statistics is that they "best honor the reality to which the researcher is purportedly trying to generalize" (Thompson, 1988b, p. 12). This reality is human behavior, which involves multiple outcomes and multiple causes (Thompson, 1988b). "These multivariate methods allow understanding of relationships among several variables not possible with univariate analysis (Hopkins, 1980, p. 374). Similarly, McMillan and Schumacher (1984) have asserted that "in the reality of complex social situations the researcher needs to examine many variables simultaneously" (p. 270), which is not possible with univariate procedures.

Since multivariate methodology is the appropriate choice in behavioral research, the question remains: which multivariate procedure is appropriate? While Fish (1988) maintains that the most frequently used multivariate methods are (MANOVA), discriminant analysis, and canonical correlation analysis, Hopkins (1980) contends that "factor analysis, canonical correlation, and discriminant analysis .. allow researchers to study complex data, particularly situations with many interrelated variables" (p. 374).

OVA methods have been used extensively, especially in dissertation research; however, as Thompson (1986a) has shown, general linear models, such as multiple regression and canonical correlation analysis, are superior. The main objection to OVA techniques is that they "require that all independent variables be nominally scaled" (Thompson, 1986a, p. 918). As Thompson (1981, p. 8) has stated, "When we reduce interval level of scale to the nominal level of scale we are doing nothing less than thoughtlessly throwing away information which we previously went to some trouble to collect." Other objections to OVA methods include the distortion of distribution shapes and relationships among variables caused by the forced but computationally simple balanced design preferred by many OVA researchers (Thompson, 1986a), and the reduction of reliability and consequent reduction of power against Type II errors caused by the conversion of intervally scaled data to the nominal level (Cohen, 1968).

Canonical correlation analysis is the most general case of the general linear model (Baggalley, 1981). Knapp (1978) demonstrated that canonical correlation analysis subsumes all other parametric tests as special cases. Thompson (1988a) used a small hypothetical data set to illustrate how canonical correlation gives the same results as t tests, Pearson product moment correlations, ANOVA, MANOVA, multiple regression, and discriminant analysis.

Canonical correlation analyses "create synthetic composite variable sets for both the independent variables and the dependent variables" (Campbell, 1990a, pp. 89-90). A canonical analysis is simply the Pearson product moment correlation between the composite set of independent (predictor) variables and the composite set of dependent (criterion) variables.

Since canonical correlation analysis subsumes multiple regression as a special case, and since commonality analysis has proven useful in interpreting multiple regression results (Thompson & Borello, 1985), then the interpretation of canonical correlation results might also be enhanced by the use of commonality analysis. Commonality analysis is "an attempt to understand the relative predictive power of the regressor [independent] variables, both individually and in combination" (Beaton, 1973, p. 2). As Daniel (1989, p. 4) points out, by using commonality analysis, "a researcher can determine the unique and the common contributions of each independent variable and each interaction effect in a prediction equation."

In the present paper a canonical correlation analysis using a data set of 64 observations, consisting of four predictor (independent) variables and four criterion (dependent) variables, is interpreted. A commonality analysis of the data is then explained and illustrated. These results are then interpreted in conjunction with the canonical correlation analysis.

The data set involved 64 subjects or observations. The predictor variables were raw scores on four scores of the Myers-Briggs Type Indicator (MBTI) (Briggs & Myers, 1976), an instrument which measures the psychological dimensions espoused by Carl Jung's theory of psychological type. The criterion variables were the scores on the Information Preference Questionnaire (IPQ) (Campbell, 1990a), an instrument designed by the author to measure subjects' self-reported preferences for certain types of information.

The MBTI scores used as independent variables in the present study reflected both the manner in which individuals perceive information--Sensing or Intuiting-- and the manner in which individuals make decisions--Thinking or Feeling. "The crux of Jung's theory is that individuals perceive information either through sensing or through intuition, and they make judgments on their perceived information either through thinking or through feeling" (Campbell, 1991, p. 3).

The resultant combinations of individuals' preferred manners of perception and judgments produce four cognitive styles: Sensing Thinking (ST), Sensing Feeling (SF), Intuitive Thinking (NT), and Intuitive Feeling (NF). "Current cognitive style research indicates that cognitive style influences the kind of information a decision maker values as important" (Campbell, 1991, p. 5).

The IPQ scores used as dependent variables in the present study reflected the preferences of individuals among four types of information constructed by the author to match the four Jungian cognitive styles (ST, SF, NT, NF). These four types of information were presented after each of 12 scenarios concerning school related decisions.

A canonical correlation analysis was performed on the data. The shared variance of the MBTI scores and IPQ scores was measured by the squared canonical correlation. The squared canonical correlation indicates the effect size, "the measure of the percentage of the variance of the dependent variables that is shared with the independent variables" (Campbell, 1991, p. 11-12). Table 1 presents the canonical results. The squared canonical correlation of .212 indicates that 21% of the variance of information preference scores can be explained by MBTI scores. In other words, he MBTI scores can predict 21% of the variation of information preferences among school principals.

[Insert Table 1 here.]

Commonality analysis, according to Thompson (1988a), facilitates interpretation of canonical results. Thompson and Miller (1985, p. 2) maintain that the "analysis indicates how much of the explanatory power of a variable is 'unique' to the variable, and how much of the variable's explanatory ability is 'common' to or also available from one or more other variables." Commonality analyses may be performed in univariate multiple regression cases or in multivariate canonical correlation analyses. The only difference between the univariate and the multivariate analysis is that in the multivariate case the dependent variables must be converted to a composite score, the synthetic composite score (Daniel, 1989).

Campbell (1990b) lists the steps for performing a commonality analysis: (1) perform a canonical correlation analysis; (2) calculate the z-scores and the criterion (dependent variable) composite scores; (3) calculate the regression equations, using all possible combinations of predictors (independent variables) to predict the synthetic criterion (dependent variable) composite scores); (4) calculate the unique and common variance effects, and then add the columns for each predictor (independent) variable to find the sum of the explanatory power of each predictor. In the present case, a commonality analysis was conducted only on the first canonical function; however, a commonality analysis could also have been conducted on the other canonical functions.

The z-scores for the dependent variables were calculated using the formula, the score minus the mean divided by the standard deviation. Once the z-scores were calculated, the dependent variables were put into a synthetic composite set "by summing the products of the standardized function coefficients and the z-scores of the dependent variables" (Campbell, 1990a, p. 132). The standardized function coefficients are actually beta weights and are used to multiply the z-scores of the dependent variables, after which these products are summed to create the composite set of criterion (dependent) variables.

After creating the composite set of dependent variables, a multiple regression was run on the composite dependent variable set, using all possible combinations of the independent variables. (This may be accomplished through the PROC RSQUARE command in the SAS program.) Table 2 shows the coefficient Rc2 for all possible regression combinations on the first canonical function. The Rc2, when all four predictors (independent variables) are used in combination, equals the squared canonical correlation.

[Insert Table 2 here.]

After the regression equations are calculated, the unique and common variance partitions must be calculated. Thompson and Miller (1985) provided the formula for a set of four independent variables, which is the number in the present study. Table 3 summarizes the partitioning calculations. Negative calculations do not indicate that these variables have a less than zero effect but, rather "the presence of suppressor effects, such as negative correlations between variables" (Campbell, 1990b, p. 9).

[Insert Table 3 here.]

After calculating the unique and common variances, those results were entered into the appropriate partitions for each of the independent variables, and the columns for each independent variable were summed to determine the explanatory power of each independent variable alone. The results of the commonality analysis are summarized in Table 4. In the present case the MBTI cognitive functions Thinking and Feeling had the greatest explanatory ability.

[Insert Table 4 here.]

Summary

Commonality analysis is a variance partitioning procedure which facilitates the interpretation of multiple regression and canonical correlation analysis results. Daniel (1989) and Thompson and Miller (1985) have pointed several advantages to the procedure. First, commonality honors the relationships among variables by indicating the degree to which predictors (independent variables) in a set share variance with the criterion (dependent) variable(s). Second, since commonality indicates the extent of overlap of the variables, it is useful in the behavioral sciences where independent variables are usually correlated with each other. Third, commonality reinforces the recognition that canonical correlation analysis is the most general linear model of parametric statistics.

REFERENCES

Baggalley, A.R. (1981). Multivariate analysis: An introduction for consumers of behavioral research. Evaluation Review, 5, 123-131.

Beaton, A.E. (1973). Commonality. Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED 111 829)

Briggs, K.C., and Myers, I.B. (1976). Myers-Briggs Type Indicator: Form G. Palo Alto, CA: Consulting Psychologists Press.

Campbell, K.T. (1990a). The influence of Jungian cognitive style on information preferences among school principals. (Doctoral dissertation, University of New Orleans, 1991).

Campbell, K.T. (1990b). Methods of multivariate commonality analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX.

Campbell, K.T. (1991). The influence of cognitive style on information preferences among school principals. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

Daniel, L.G. (1988). Commonality analysis with various parametric methods. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Fish, L.J. (1988). Why multivariate methods are usually vital. Measurement and Evaluation in Counseling and Development, 21, 130-137.

Hopkins, C.D. (1980). Understanding educational research: An inquiry approach. New York: Holt, Rinehart, and Winston.

Huberty, C.J., & Morris, J.D. (1989). Multivariate analysis versus multiple univariate analysis. Psychological Bulletin, 105, 302-308.

LeCluyse, K. (1990). Why multivariate statistics are critical: A review with examples. Paper presented at the annual meeting of the Southwest Educationsl Research Association, Austin, TX.

McMillan, J.H. & Schumacher, S. (1984). Research in education: A conceptual approach. Boston: Little, Brown.

Thompson, B. (1981). The problem of OVAism. Paper presented at the annual meeting of the Mid-South Educational Research Association, Lexington, KY.

Thompson, B. (1986a). ANOVA versus regression analysis of ATI designs: An empirical investigation. Educational and Psychological Measurement, 46, 917-928.

Thompson, B. (1986b). Two reasons why multivariate statistics are usually vital. Paper presented at the annual meeting of the Mid-South Educational Research Association, Memphis, TN.

Thompson, B. (1988a). Canonical correlation analysis: An explanation with comments on correct practice. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Thompson, B. (1988b). Common methodology mistakes in dissertations: improving dissertation quality. Paper presented at the annual meeting of the Mid-South Educational Research Association, Louisville, KY.

Thompson, B., & Borrello, G. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

Thompson, B. and Miller, J. (1985). A multivariate method of commonality analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX.

TABLE 1

Canonical Correlation Analysis Coefficients

on the First Two Canonical Functions

Total Group of Principals and Assistant Principals (n = 64)

Eliminating the Eighth Scenario of the IPQ

Var./ Coeff.	I Func	Stru	Sq Stru	II Func	Stru	Sq Stru	h^2
S	1.256	-.308	.09	2.067	.847	.72	.81
N	1.456	.503	.25	1.339	-.572	.33	.58
T	-.985	-.826	.68	.074	.280	.08	.76
F	-.217	.734	.54	.033	-.203	.04	.58
Rc^2	.212029			.103000			
ST	-1.017	-.208	.04	.014	.447	.20	.24
SF	1.152	.647	.42	.200	.338	.11	.53
NT	.147	.169	.03	1.392	.514	.26	.29
NF	.065	.276	.08	-1.309	-.161	.03	.11

Note. The h^2 values reported in the table are based on only two of the four functions, e.g., .81 = .09 + .72.

TABLE 2

Prediction of Composite IPQ Scores

Using Alternate Predictor Variable Combinations

Total Group of Principals and Assistant Principals (n = 64)

Eliminating the Eighth Scenario of the IPQ

Predictor Set	Variables in Set	Rc^2
S	ZS	.020155
N	ZN	.053704
T	ZT	.144690
F	ZF	.114108
SN	ZS, ZN	.086504
SF	ZS, ZF	.116634
NF	ZN, ZF	.132794
ST	ZS, ZT	.1457956
TF	ZT, ZF	.144790
NT	ZN, ZT	.162539
STF	ZS, ZT, ZF	.145864
NTF	ZN, ZT, ZF	.162604
SNF	ZS, ZN, ZF	.165980
SNT	ZS, ZN, ZT	.209802
SNTF	ZS, ZN, ZT, ZF	.212029

The Rc^2 for predictor set SNTF (using all four predictor variables in combination) equals the squared canonical correlation, within rounding error.

TABLE 3

Calculations of Unique and Common Variance Partitions

Total Group of Principals and Assistant Principals (n = 64)

Eliminating the Eighth Scenario of the IPQ

Set	Partition	Result
Unique to ZS	SNTF - NTF	.049424
Unique to ZN	SNTF - STF	.066165
Unique to ZT	SNTF - SNF	.046049
Unique to ZF	SNTF - SNT	.002226
Common to ZS, ZN	STF + NTF - TF - SNTF	-.048350
Common to ZS, ZT	SNF + NTF - NF - SNTF	-.016238
Common to ZS, ZF	SNT + NTF - NT - SNTF	-.002161
Common to ZN, ZT	SNF + STF - SF - SNTF	-.016819
Common to ZN, ZF	SNT + STF - ST - SNTF	-.002158
Common to ZT, ZF	SNT + SNF - SN - SNTF	.077249
Common to ZS, ZN, ZT	TF + NF + SF + SNTF - F	
	- NTF - STF - SNF	.017690
Common to ZS, ZN, ZF	TF + NF + ST + SNTF - T	
	-NTF - STF - SNT	-.027553
Common to ZS, ZT, ZF	NF + NT + SN + SNTF - N	
	- NTF - SNF - SNT	.001775
Common to ZN, ZT, ZF	SF + ST + SN + SNTF -S	
	- STF - SNF - SNT	.019162
Common to ZS, ZN, ZT, ZF	F + T + N + S + NTF + STF	
	+ SNF + SNT - TF - NF -NT	
	- SF - ST - SN - SNTF	.015822

TABLE 4

Multivariate Commonality Results

Total Group of Principals and Assistant Principals (n = 64)

Eliminating the Eighth Scenario of the IPQ

Partition	ZS	ZN	ZT	ZF
Unique to ZS	.049424			
Unique to ZN		.066165		
Unique to ZT			.046049	
Unique to ZF				.002226
Common to ZS,ZN	-.048350	-.048350		
Common to ZS,ZT	-.016238		-.016238	
Common to ZS,ZF	-.002161			-.002161
Common to ZN,ZT		-.016819	-.016819	
Common to ZN,ZF		-.002158		-.002158
Common to ZT,ZF			.077249	.077249
Common to ZS,ZN,ZT	.017690	.017690	.017690	
Common to ZS,ZN,ZF	-.027553	-.027553		-.027553
Common to ZS,ZT,ZF	.001775		.001775	.001775
Common to ZN,ZT,ZF		.019162	.019162	.019162
Common to ZS,ZN,ZT,ZF	.015822	.015822	.015822	.015822
Sum of the Partitions	-.009591	.023959	.144690	.084362
r^2 of Predictor with canonical composite of cognitive style IPQ scores	.96%	2.40%	14.47%	8.44%

Note. The underlined commonality coefficients should be large because these variable pairs are parts of a given MBTI scale.